JOURNAL OF ANIMAL SCIENCE
The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

Technical note: Validation of a model for online classification of US Select beef carcasses for longissimus tenderness using visible and near-infrared reflectance spectroscopy

S. D. Shackelford, T. L. Wheeler and M. Koohmaraie

J ANIM SCI 2012, 90:973-977.
doi: 10.2527/jas.2011-4166 originally published online November 7, 2011

Technical note: Validation of a model for online classification of US Select beef carcasses for longissimus tenderness using visible and near-infrared reflectance spectroscopy[1,2,3]

S. D. Shackelford,[4] T. L. Wheeler, and M. Koohmaraie[5]

USDA-ARS, Roman L. Hruska US Meat Animal Research Center, Clay Center, NE 68933-0166

ABSTRACT: The present experiment was conducted to provide a validation of a previously developed model for online classification of US Select carcasses for LM tenderness based on visible and near-infrared (VISNIR) spectroscopy and to determine if the accuracy of VISNIR-based tenderness classification could be enhanced by making measurements after postmortem aging. Spectroscopy was conducted online, during carcass grading, at a large-scale commercial fed beef-processing facility, and the strip loin was obtained from the left side of US Select carcasses (n = 467). Slice shear force (SSF) was measured on fresh steaks at 2 and 14 d postmortem. Online VISNIR tenderness classes differed in mean SSF values at both 2 d (29.4 vs. 33.6 kg) and 14 d (18.0 vs. 21.2 kg) postmortem ($P < 10^{-7}$). Online VISNIR tenderness classes differed in both the percentage of carcasses with LM SSF values greater than 40 kg at 2 d postmortem (5.1 vs. 21.0%; $P < 10^{-6}$) and the percentage of carcasses with LM SSF values greater than 25 kg at 14 d postmortem (6.8 vs. 23.2%; $P < 10^{-5}$). Whereas 15.0% of the carcasses sampled for this experiment had LM SSF values greater than 25 kg at 14 d postmortem, only 6.8% of the carcasses classified as tender by VISNIR had LM SSF values greater than 25 kg. All the carcasses sampled that had LM SSF values greater than 35 kg at 14 d postmortem were accurately classified as tough by VISNIR. Before measurement of SSF on d 14, VISNIR spectroscopy was conducted on the SSF steak. Tenderness classes based on d 14 VISNIR spectra differed both in mean SSF value at 14 d postmortem (17.7 vs. 21.6 kg; $P < 10^{-11}$) and the percentage of carcasses with LM SSF values greater than 25 kg at 14 d postmortem (7.3 vs. 22.7%; $P < 10^{-5}$). These data support our previous work showing that VISNIR spectroscopy can be used to classify US Select carcasses noninvasively for LM tenderness, and the results establish that this technology could also be applied to aged US Select strip loins. This technology would allow packing companies and other segments of the beef marketing chain to identify US Select carcasses or strip loins that excel in LM tenderness for use in branded beef programs.

Key words: beef, near-infrared, prediction, slice shear force, tenderness

J. Anim. Sci. 2012. 90:973–977
http://dx.doi.org/10.2527/jas.2011-4166

INTRODUCTION

Shackelford et al. (2004) developed a highly repeatable method for online spectroscopic evaluation of LM quality traits of ribbed beef carcasses using a high-intensity reflectance probe that allowed for sampling of a representative portion of the exposed LM cross-section of ribbed beef carcasses. Shackelford et al. (2005) developed and tested models for online classification of US Select beef carcasses for LM tenderness by using visible and near-infrared (**VISNIR**) reflectance spectroscopy. Although not highly accurate, VISNIR-based sorting allowed segmentation of the carcasses into 2 groups that differed ($P < 0.001$) in mean slice shear force (**SSF**; 16.3 vs. 21.6 kg) and the proportion of carcasses with SSF values greater than 25 kg (5.5% vs. 30.1%). The present experiments were conducted to 1) provide a validation of the previously developed model

[1]This project was funded, in part, by The Beef Checkoff (Denver, CO).

[2]Mention of trade names, proprietary products, or specified equipment does not constitute a guarantee or warranty by the USDA and does not imply approval to the exclusion of other products that may be suitable. The authors are grateful to Patty Beska, Ryan Dijkhuis, Andy Everts, Lauren Grimes, Beryl Henry, Kathy Mihm, Pat Tammen, and Andrea Warfield of the US Meat Animal Research Center for their assistance in the execution of this experiment and to Marilyn Bierman of the US Meat Animal Research Center for her secretarial assistance.

[3]USDA is an equal opportunity provider and employer.

[4]Corresponding author: steven.shackelford@ars.usda.gov

[5]Current address: IEH Laboratories and Consulting Group, 15300 Bothell Way NE, Lake Forest Park, WA 98155.

Received April 12, 2011.
Accepted October 27, 2011.

for online classification of US Select carcasses for LM tenderness both before and after postmortem aging, and 2) to determine if the accuracy of VISNIR-based tenderness classification could be enhanced by making measurements after postmortem aging.

MATERIALS AND METHODS

Animal procedures were not reviewed and approved by the US Meat Animal Research Center (**USMARC**) Animal Care and Use Committee because this experiment did not involve animals originating from or under the control of USMARC.

Spectroscopy and Carcass Selection

Spectroscopy was conducted online at a large-scale commercial fed beef-processing facility. This facility was 1 of the 2 plants in which Shackelford et al. (2005) collected spectra to develop the tenderness prediction model. To facilitate data collection and to ensure broadness of sampling, testing was conducted 1 d per week for 3 consecutive weeks with 134, 168, and 165 samples obtained on the 3 respective selection trips. At approximately 26 h postmortem, carcasses were ribbed conventionally between the 12th and 13th ribs for determination of USDA beef quality and yield grades, and spectroscopy was conducted on the LM cross section. Spectroscopy was conducted on the beef grading bloom chain approximately 2 min after the carcasses were ribbed. Spectroscopy was conducted only on carcasses that were likely to be graded USDA Select and USDA yield grade 1, 2, or 3. Spectroscopy was conducted on the left side of each carcass by using the optimal protocol (a single measurement averaging 20 spectra per measurement) and the spectroscopy system described by Shackelford et al. (2004), and SSF was predicted using the model described by Shackelford et al. (2005). After grading, any carcasses that did not meet these specifications were excluded from the experiment.

After spectroscopy and grading, carcasses were fabricated and the NAMP 180 strip loin (North American Meat Processors Association, 2010) was obtained from the left side of each carcass (n = 467). Strip loins were transported to the USMARC abattoir and held overnight (1°C), and then each loin was trimmed free of subcutaneous fat. A 2.54-cm-thick steak was removed from the anterior end of each strip loin and cooked from the fresh (2 d postmortem; never frozen) state using a belt grill (Wheeler et al., 1998), and SSF was measured (Shackelford et al., 1999a,b). The remainder of each strip loin was vacuum-packaged and aged (1°C) until 14 d postmortem. Again, a 2.54-cm-thick steak was removed from the anterior end of each strip loin and VISNIR spectroscopy was conducted on the fresh cut face of the steak approximately 10 s after the steak was cut, using the same procedures as above, and then the steak was cooked using a belt grill and SSF was measured.

Statistical Analysis

The data (NIR spectra and observed SSF) were not checked for outliers. Thus, selective culling of the data, which could affect the results and the utility of the prediction models, was not conducted.

Model Validation. Slice shear force was predicted online in the packing plant by using the model developed by Shackelford et al. (2005). To facilitate statistical analysis, the data were divided into halves. Specifically, those carcasses with VISNIR-predicted SSF values less than or equal to the median were classified as "VISNIR-Predicted Tender" and those carcasses with VISNIR-predicted SSF values greater than the median were classified as "VISNIR-Not Predicted Tender." One-way ANOVA for SSF at both 2 and 14 d postmortem was conducted using the GLM procedure (SAS Inst. Inc., Cary, NC) for the single classification variable of the VISNIR-predicted tenderness class. The frequency of carcasses with SSF values >25 kg was calculated for each VISNIR class. Differences in these frequencies among VISNIR classes were compared using the DIFFER program (PEPI, USD Inc., Stone Mountain, GA.).

d 14 Prediction. Prediction of SSF with d 14 spectra was conducted in 2 ways. First, SSF was predicted using the model developed by Shackelford et al. (2005). Second, a model was developed using the PLS1 procedure (The Unscrambler v9.8, CAMO Software AS, Oslo, Norway). In both cases, spectra were reduced by averaging groups of 9 consecutive wavelengths. For model development and validation, carcasses were blocked by observed SSF. One-half of the carcasses were assigned to a calibration data set, which was used to develop regression equations, and one-half of the carcasses were assigned to a prediction data set, which was used to validate the regression equations (Neter et al., 1989). Specifically, carcasses were ranked by SSF from least to greatest, and alternating carcasses were assigned to the calibration and prediction data sets. Thus, all simple statistics of observed SSF were virtually identical for the calibration (mean = 19.6 kg; SD = 6.2 kg; minimum = 8.2 kg; and maximum = 45.9 kg) and prediction (mean = 19.6 kg; SD = 6.1 kg; minimum = 9.0 kg; and maximum = 45.8 kg) data sets. All analyses of efficacy of tenderness sorting were conducted using the prediction data set.

RESULTS AND DISCUSSION

Model Validation

Online VISNIR tenderness classes differed in mean SSF values at both 2 and 14 d postmortem (Figure 1; $P < 10^{-7}$). Online VISNIR tenderness classes differed in both the percentage of carcasses with LM SSF values greater than 40 kg at 2 d postmortem (5.1 vs. 21.0%; $P < 10^{-6}$) and the percentage of carcasses with LM SSF values greater than 25 kg at 14 d postmortem (6.8 vs. 23.2%; $P < 10^{-5}$). Whereas 15.0% of the carcasses sampled for this experiment had LM SSF val-

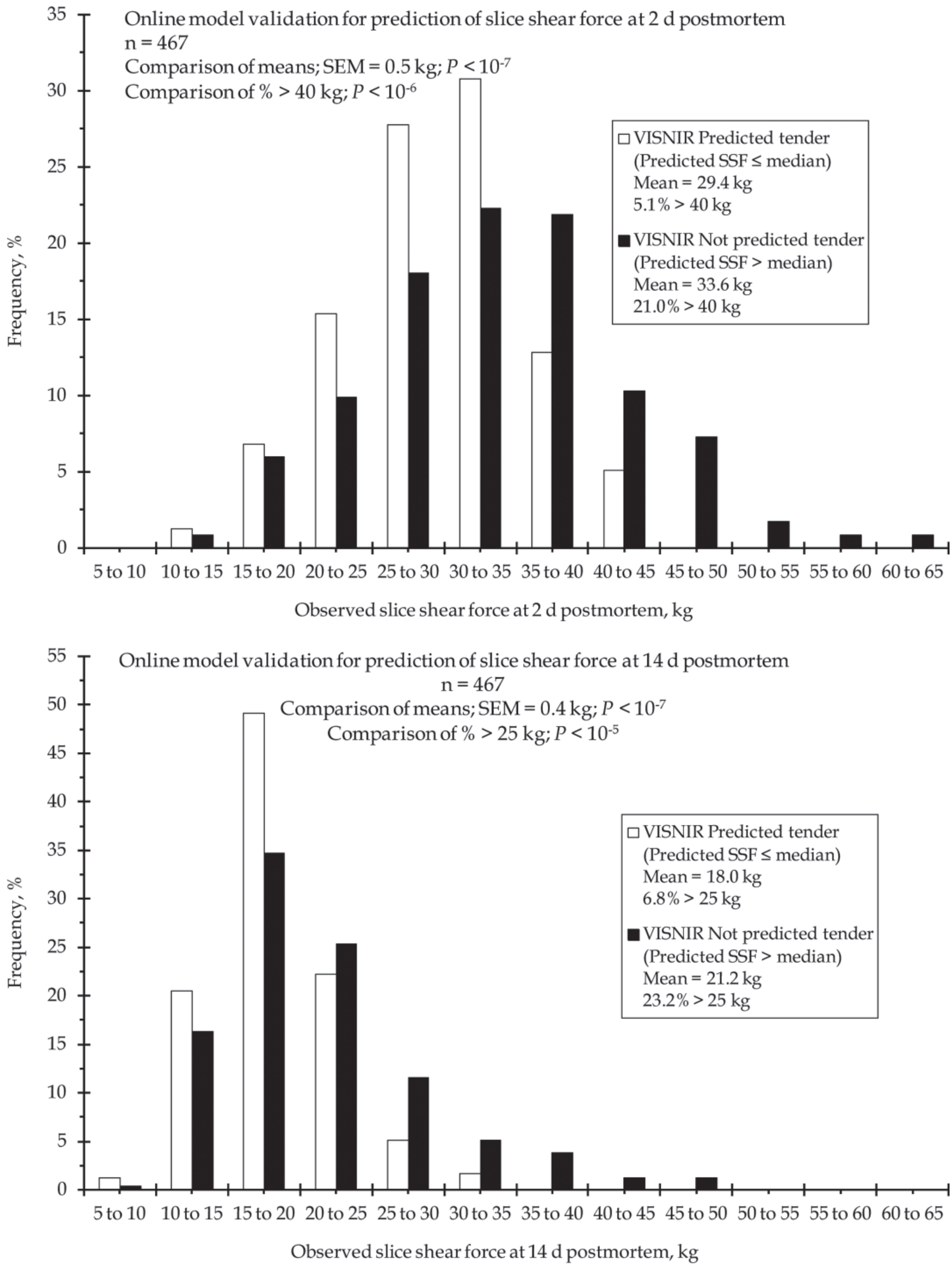
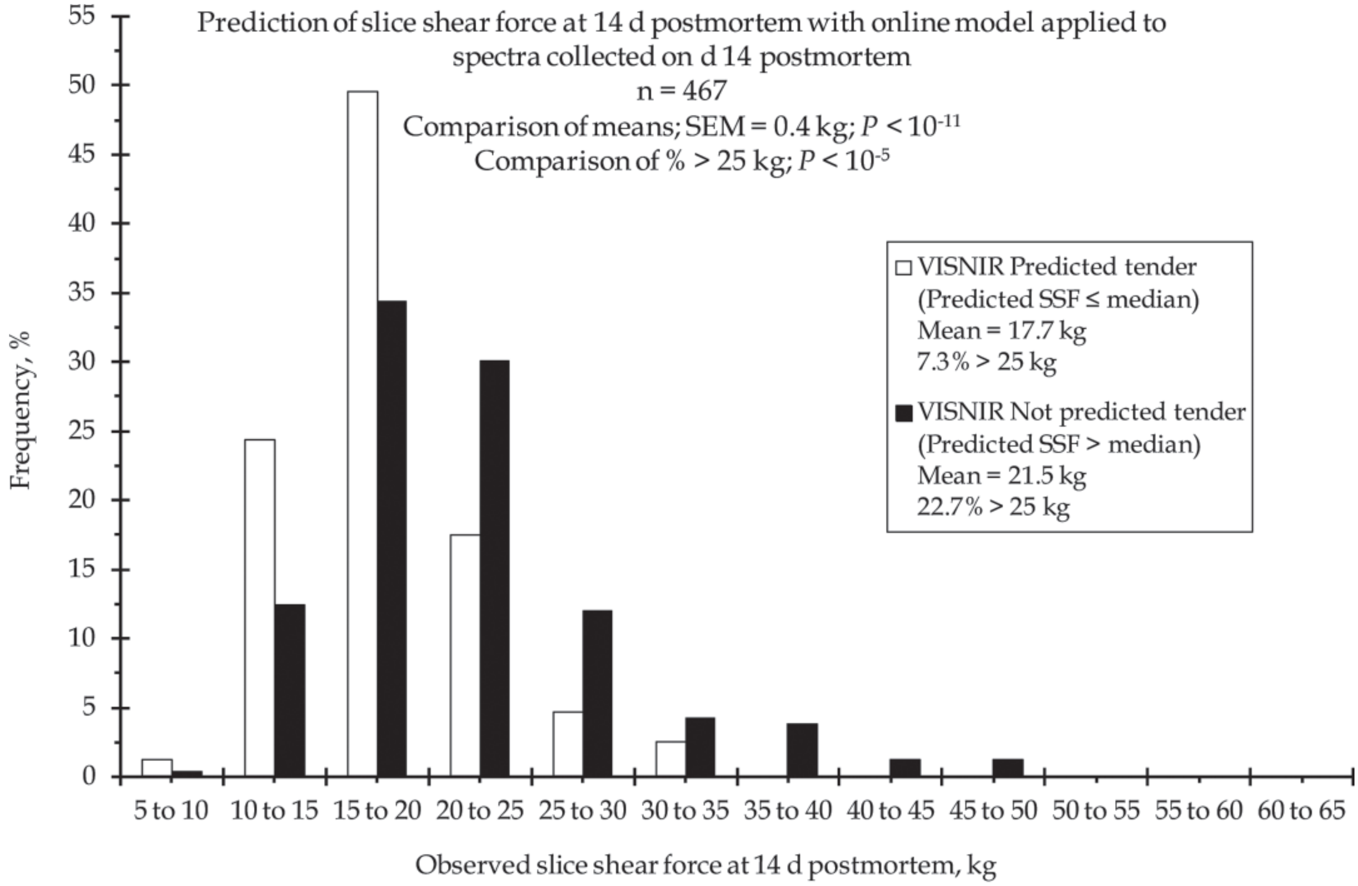


Figure 1. Effect of online sorting of carcasses into predicted tenderness classes on LM slice shear force (SSF) at 2 d (top panel) and 14 d (bottom panel) postmortem. Sorting was based on online evaluation of the ribeye with visible and near-infrared (VISNIR) spectroscopy by using the prediction model of Shackelford et al. (2005).



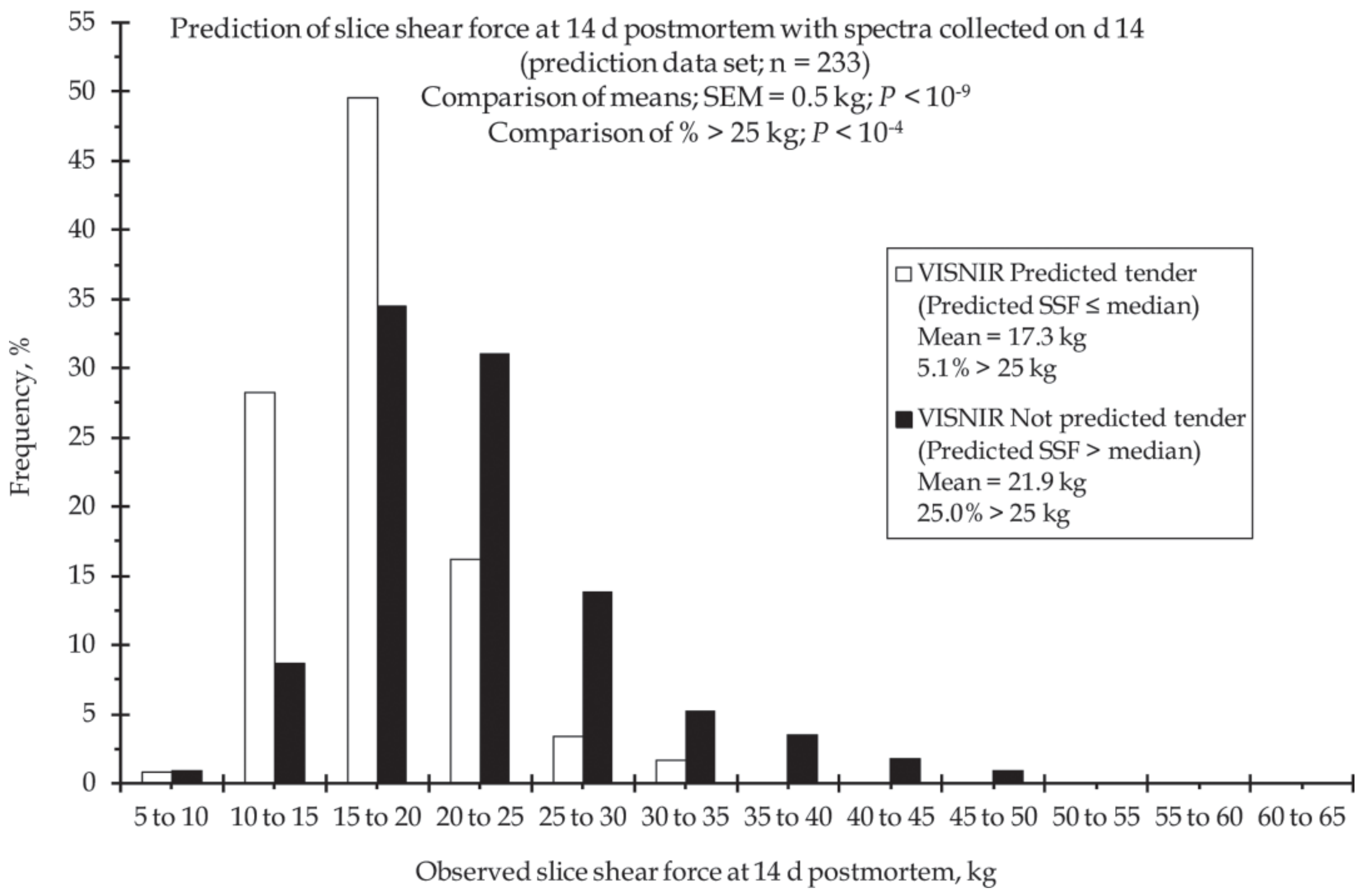


Figure 2. Effect of sorting strip loins into predicted tenderness classes on LM slice shear force (SSF) at 14 d postmortem. Sorting was based on evaluation of an LM steak with visible and near-infrared (VISNIR) spectroscopy by using either the prediction model (top panel) of Shackelford et al. (2005) or a model calibrated in the present experiment using spectra collected at 14 d postmortem (bottom panel).

ues greater than 25 kg at 14 d postmortem, only 6.8% of the carcasses classified as tender by VISNIR had LM SSF values greater than 25 kg. All the carcasses sampled that had LM SSF values greater than 35 kg at 14 d postmortem were accurately classified as tough by VISNIR (Figure 1). These data support our previous work and that of Price et al. (2008) showing that VISNIR spectroscopy can be used online to classify US Select carcasses noninvasively for LM tenderness. This technology would allow packing companies to identify US Select carcasses that excel in LM tenderness for use in branded beef programs.

Application of VISNIR to Aged Cuts

Although the results described above showed that the use of VISNIR is efficacious for tenderness sorting at the time of grading and sorting, it was not clear what the potential usefulness of this technology might be to other sectors of the industry. For instance, could a grocery chain apply this technology to aged strip loins as part of a centralized steak-cutting operation and identify more tender cuts to go to a premium brand and less tender cuts to go to a generic brand? When VISNIR spectra were collected at 14 d postmortem and SSF was predicted using the model developed by Shackelford et al. (2005), VISNIR tenderness classes differed in both mean SSF value at 14 d postmortem ($P < 10^{-11}$) and the percentage of carcasses with LM SSF values greater than 25 kg at 14 d postmortem ($P < 10^{-5}$; top panel of Figure 2). Similarly, when VISNIR spectra that were collected at 14 d postmortem were used to develop (one-half of the observations) and validate (the other one-half of the observations; $R^2 = 0.26$; root mean square error = 5.2 kg) a model to predict SSF, VISNIR tenderness classes differed both in mean SSF value at 14 d postmortem ($P < 10^{-9}$) and the percentage of carcasses with LM SSF values greater than 25 kg at 14 d postmortem ($P < 10^{-4}$; bottom panel of Figure 2).

These data indicate not only that it would be possible to apply this technology to aged LM cuts, but also that it could be done using the same model as was used for online evaluation of carcasses (i.e., unaged) with very little loss of accuracy relative to using a model developed on aged cuts. Although it would be desirable to have tenderness predicted in each scenario with as much precision as possible, it would make it much simpler to implement VISNIR commercially if a single model could be used for unaged and aged cuts. This would be particularly true at the packing plant level, where the time postmortem at which carcasses are graded varies greatly within and among plants.

The fact that the same model could successfully predict the tenderness of unaged and aged beef indicates that the system is not directly assessing postmortem proteolysis. This does not exclude the possibility that the system is assessing a trait that affects postmortem proteolysis. Furthermore, the prediction could be completely unrelated to postmortem proteolysis.

General Discussion

These data both confirm the findings of Shackelford et al. (2005) and establish that this technology for noninvasive tenderness prediction can be robust across time when applied under a similar situation (i.e., the same packing plant). These results provide the justification for studies to establish the robustness of this technology across additional plants, geographical regions, and cattle or carcass types. This study was intentionally focused on a very narrow application to US Select carcasses because of the potential economic rewards associated with identifying consistently tender Select carcasses (Shackelford et al., 2001). Additional work is needed to establish efficacy for other quality grades. The results of Rust et al. (2008) seem to indicate that a similar, but off-line, system was efficacious across quality grades (i.e., a pool of US Select and US Choice carcasses); however, it is unclear whether their system was efficacious within US Choice.

The present experiment confirms that US Select carcasses can be noninvasively classified online for LM tenderness by using VISNIR spectroscopy, and it establishes that this technology could be applied to steaks obtained from aged US Select strip loins. This technology could facilitate tenderness-based beef merchandising systems.

LITERATURE CITED

Neter, J., W. Wasserman, and M. H. Kutner. 1989. Applied Linear Regression Models. 2nd ed. Irwin, Homewood, IL.

North American Meat Processors Association. 2010. Meat Buyers Guide. North Am. Meat Process. Assoc., Reston, VA.

Price, D. M., G. G. Hilton, D. L. VanOverbeke, and J. B. Morgan. 2008. Using the near-infrared system to sort various beef middle and end cuts into tenderness categories. J. Anim. Sci. 86:413–418.

Rust, S. R., D. M. Price, J. Subbiah, G. Kranzler, G. G. Hilton, D. L. Vanoverbeke, and J. B. Morgan. 2008. Predicting beef tenderness using near-infrared spectroscopy. J. Anim. Sci. 86:211–219.

Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999a. Evaluation of slice shear force as an objective method of assessing beef longissimus tenderness. J. Anim. Sci. 77:2693–2699.

Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 1999b. Tenderness classification of beef. II: Design and analysis of a system to measure beef longissimus shear force under commercial processing conditions. J. Anim. Sci. 77:1474–1481.

Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 2004. Development of optimal protocol for visible and near-infrared reflectance spectroscopic evaluation of meat quality. Meat Sci. 68:371–381.

Shackelford, S. D., T. L. Wheeler, and M. Koohmaraie. 2005. Online classification of US Select beef carcasses for longissimus tenderness using visible and near-infrared reflectance spectroscopy. Meat Sci. 69:409–415.

Shackelford, S. D., T. L. Wheeler, M. K. Meade, J. O. Reagan, B. L. Byrnes, and M. Koohmaraie. 2001. Consumer impressions of Tender Select beef. J. Anim. Sci. 79:2605–2614.

Wheeler, T. L., S. D. Shackelford, and M. Koohmaraie. 1998. Cooking and palatability traits of beef longissimus steaks cooked with a belt grill or an open hearth electric broiler. J. Anim. Sci. 76:2805–2810.